OIPE

RAW SEQUENCE LISTING DATE: 07/05/2001
PATENT APPLICATION: US/09/885,297 TIME: 16:29:04

Input Set : A:\seqlist24CP.txt

Output Set: N:\CRF3\07032001\1885297.raw

Does Not Comply
Corrected Diskette Needed

pr1,6-9

3 <110> APPLICANT: Ingram, L et al.

5 <120> TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SIMULTANEOUS

6 SACCHARIFICATION AND FERMENTATION

8 <130> FILE REFERENCE: BCI-024CP

C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/885,297

C--> 11 <141> CURRENT FILING DATE: 2001-06-19

13 <150> PRIOR APPLICATION NUMBER: 60/214,137

14 <151> PRIOR FILING DATE: 2000-06-26

16 <150> PRIOR APPLICATION NUMBER: 60/219,913

17 <151> PRIOR FILING DATE: 2000-07-21

19 <160> NUMBER OF SEQ ID NOS: 17

21 <170> SOFTWARE: PatentIn Ver. 2.0

ERRORED SEQUENCES

- 863 <210> SEQ ID NO: 17
- 864 <211> LENGTH: 11772
- 865 <212> TYPE: DNA
- 866 <213> ORGANISM: Artificial Sequence
- 868 <220> FEATURE:
- 869 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
- 870 construct
- 872 <220> FEATURE:
- 873 <223> OTHER INFORMATION: unsequenced Erwinia DNA fragment from nucleotide
- 874 position 1143 to 144
- 876 <220> FEATURE:
- 877 <223> OTHER INFORMATION: P1 promoter region for celz from nucleotide
- 878 position 4424 to 2974
- 880 <220> FEATURE:
- 881 <223> OTHER INFORMATION: guide fragment for integration from nucleotide
- 882 position 4677 to 7573
- 884 <220> FEATURE:
- 885 <223> OTHER INFORMATION: sequenced partial guide fragment from nucleotide
- 886 position 4677 to 5752
- 888 <220> FEATURE:
- 889 <223> OTHER INFORMATION: unsequenced partial guide fragment from nucleotide

A [

- 890 position 5753 to 7573 **OK**
- 892 <220> FEATURE:
- 893 <223> OTHER INFORMATION: P2 promoter region for cell from nucleotide
- 894 position 7585 to 8576
- 896 <220> FEATURE:
- 897 <223> OTHER INFORMATION: R6K-Y ori from nucleotide position 10388 to 10763
- 899 <220> FEATURE:
- 900 <223> OTHER INFORMATION: FRTF lipase-binding sequence from nucleotide
- 901 position 16 to 50
- 903 <220> FEATURE:

to last location

RAW SEQUENCE LISTINGPATENT APPLICATION: **US/09/885,297**DATE: 07/05/2001

TIME: 16:29:04

Input Set : A:\seqlist24CP.txt

Output Set: N:\CRF3\07032001\1885297.raw

```
904 <223> OTHER INFORMATION: FRTFlipase-binding sequence from nucleotide
         position 10058 to 10092
   905
   907 <220> FEATURE:
   908 <223> OTHER INFORMATION: celZ gene product is encoded by the complement of
         nucleotides 2973 to 1690
   909
   911 <220> FEATURE:
   912 <223> OTHER INFORMATION: celY gene product is encoded by the nucleotides
   913
         8576 to 9574
   915 <220> FEATURE:
   916 <223> OTHER INFORMATION: kanamicin-resistance gene product is encoded by
         the complement of nucleotides 11621 to 10827
   919 <400> SEQUENCE: 17
   920 atcgatgaat tgatccgaag ttcctattct ctagaaagta taggaacttc gaattgtcga 60
   921 caagettgat etggettate gaaattaata egaeteaeta tagggagaee ggaatteeee 120
W--> 922 tgcaggtcga ctctagagga tcannnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn 180
W--> 939 nnnagcgcgt tcgccggcaa tgccaacggg ttcaaactgg gaggaaccag cggtcggcaa 1200
   940 tcatcgcatt acccgttgcg tcgcgtttgg caacgtgagc aaaggttttg accagaacaa 1260
   941 taacgcggcg gcgtacggta ataaataaca cgtcatataa aaacggcatt aattatgctt 1320
   942 togggagoaa cgtocagtoa ggacagaaac attacttoog caacaacgtg tototgtotg 1380
   943 cttctgcaac ggtcagtaat gcggatgcga gtcaactcat ggatacggac agcggcttcc 1440
   944 gcgtccgatt cgtcagcctg gacacctcgc tggctaccgt atctcgtgat aatgacggca 1500
   945 cgttgccggc aacctcgctg ttccgtctgt cggctaactc aaagctgatt aacgcgggta 1560
   946 cgaaagagaa atcatattag ttacaaggca gcgcaccgga tttggggcgct tttgaacgca 1620
   947 attaagtaat ggaatgttgc gtatcaggct gcaatcgcag ccctgttatt ttgggggtga 1680
   948 aaagattaat caattagtta cagctaccaa cctgtgccca ggaggaatcg ctgcccggaa 1740
   949 cggatgaagt gtaccagttt gcggtataca ggttcccttt gtagacgatc gattggcctg 1800
   950 cttcqttatq aqtcqctqcc qqcccqccca gtctttqcta acccagttqq gqtaaacgtt 1860
   951 ggcattgcag caatcagcgg tttgccggtg tgtcggttgt ggtcggttcg tcaacggtgg 1920
   952 tgccatggtg gtatcggttg actgatcggt tgttgtactg gcggcgctgc ccgctttata 1980
   953 tggccagctt tgaatgatcg attttactat tttacccgac tcggtcaggt ttttagagtc 2040
   954 cggataatag gttgatgccc cttcgctttt atcatttaac gcccagtttg cgttgctgat 2100
   955 gttgttgtca cgcatgaacg ttacccaggc gtcggtatct gtctggttca ctccgccatt 2160
   956 gccgtccgcg ttaacggcgc cccactcggt gacgaaaagc gcaataccgt tatttaacgc 2220
```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/885,297

DATE: 07/05/2001 TIME: 16:29:04

Input Set : A:\seqlist24CP.txt

Output Set: N:\CRF3\07032001\I885297.raw

```
957 ctggcgggct ttagtgcgta atgactcacc atgggttccc gcgtagaaat gcagcgtata 2280
958 ggcgatatte ttggcgttga ttggatcgcg cgacgettea teaacgtttt gcgaccaact 2340
959 gggcgtaccg acaataatca ggttatccgg gtcaatggcg cgaatggcgg aaatcacggc 2400
960 ttcggcataa ggtttaatgg tattgctcca tgaaacctga agcggctcgt tgtagatttc 2460
961 ataaatgaca ttcggcttgt tgccatattt gcgcgccatt tcctggaaga agcgaatggc 2520
962 ttcactgcga ttgttttctg cagaatgtga gtgccagtca ataatcacat acatatcgtt 2580
963 ggcgattgcg gcatccacca ctctttcaac tttggccttg ttgccagccg ggtcctgcag 2640
964 ataaccaccg ctttcctgaa cgcccatagc ggcgcgaaca atgctggatt tccagtcttt 2700
965 tttcagcgac gcaacggtat cggctgtgta gaatttttcc ccaccccaac cattattact 2760
966 ccagaataag ctgttgccgg caaaactttt ggctttttca cctgcgtaga ttttattgcc 2820
967 gctaacggat aacggctcaa cactcgccca ggcattactg gaaaggcagg caaggcttaa 2880
968 tectaaacag geacaagata gaaacagttt tttaegtggg gegtgttttt ggetategat 2940
969 gactggatgg ttatccgaat aagagagagg cataaatgaa tctccatttc atagactcta 3000
970 gaggatcacg aaaatcaggc agcgcttaaa tctgccaaaa caccggaggc tgcttcatta 3060
971 tggcagcgca aggcaatttc agtataaagg gcatcacgat cgaaatcaat gctgccgccc 3120
972 tgaatggett ccatcaggte ttetgegeeg acaacttetg caccageage tttagettet 3180
973 tcagctttcg ggccacgggc gaaaacaccg acgcgaacat ttitaccggt gcccttcggc 3240
974 aaggtaacaa caccacgaac catctgatca aggtattaag gccgccggac aggtaaaatg 3300
975 gttatggcta accgatgccg atattgcgca tcaggatgat accttgaaaa tgttggttgc 3360
976 tcatgccgaa aaagacggat tgatttttca ttccctgatg gcgcgactaa gttgtaacaa 3420
977 ttttgcagaa tgggcgctta ttccggcctt cgtctttttc ttccagatgt tgtatccttt 3480
978 tcggcatatc aataatccca agcataaaac ggccggtgct gccggtggat gcatgttagc 3540
979 cgatcatcat ctgtgctggt ttggataaac cgttactaat ctggcgcgcg acccgataga 3600
980 cagcggcttc cccagccccc agtgagctac caatagcaag ggtgccaccc tgctggaaaa 3660
981 tggagtctag gatttcatta aaactgacgc ttaaagtgaa tttccacatc ccttcaacgg 3720
982 gcgtttttct ctgaaatatt tttcttatcg gaaatgctcg ctccgtatat tggtgaaaga 3780
983 gataaatgcc agcgtaacta caggtgacaa aaagcgtgaa ttgcgtcagg caccatataa 3840
984 acaaaaata ccccaaaggc atatgaagcc aataaccaat gccacaacct ccggttctga 3900
985 cgcaggtcgt aataaactca taaatagtga ccagtttaaa gcggttacac agccgcaaca 3960
986 tcccggttgg accagccgat attcataaaa agtataatca gcatacaaag caaggcatct 4020
987 ggcttgacct cggcaggcca tcccaatctt gaagtgccta aaatcaagat accgcctaac 4080
988 cctaccaaca tacctatcgc cgcactaaaa aaatcggctc tgatgcaaaa ggcaaggaca 4140
989 tcatcaaatt ggttaaaatc tttttcctga aaagcttttg aaccgtaatg cagcaaggtt 4200
990 tgccatgact gaaaacgggt aatatccgaa atcagggtag caaaagtcgt aatcaaaagc 4260
991 atgacaccaa acaggttcaa tcccaatgtc tgggatgtcc atgcgacata gacaaaactg 4320
992 cataccgcat tcaaaacacg gcctgtaatc agaatgccgg tattgccaat gatacgggaa 4380
993 aagacagatc gaatttgatg ccgtggataa attgccggtt gatcgatccc cgggtaccga 4440
994 getegaatte egagettgge gegeetatge ggtgtgaaat acegeacaga tgegtaagga 4500
995 gaaaataccq catcaggcgc cattcgccat tcaggctgcg caactgttgg gaagggcgat 4560
996 cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat 4620
997 taagttgggt aacgccaggg ttttcccagt cacgacgttg taaaacgacg gccagtgaat 4680
998 tccatcaacg cttgctgtaa ccaggagcca aagctatgaa tgtacctttt agctactcgt 4740
999 cacccaccct gagcgttgag gcgttaaagc actctattgc ttataagctg atgtttatca 4800
1000 teggeaaaga eeeggetate getaacaage atgaatgget caaegeeaeg etgttegeeg 4860
1001 ttcgcgatcg tatggttgag cgctggctgc gctcaaaccg cgcgcacgtc tctcaggaag 4920
1002 ttcgccaggt ttactacctg tcgatggaat ttttgattgg ccgtacgttg tccaacgcgc 4980
1003 tgttatcgct cggcatttat gaggatgtga acagcgcgct ggaagagatg gggctgaacc 5040
1004 ttgaagaatt aattgatgaa gaaaacgacc cgggcttagg caacggcggt cttggtcgtc 5100
1005 tggcggcctg cttcctcgat tcgcttgcgg cgctgggggtt accgggccgc ggctacggta 5160
```

RAW SEQUENCE LISTING

TIME: 16:29:04 PATENT APPLICATION: US/09/885,297

DATE: 07/05/2001

Input Set : A:\seqlist24CP.txt

Output Set: N:\CRF3\07032001\1885297.raw

| | 1006 | ttcqctacqa | ctacgggatg | tttaagcaga | atatcgtcga | tgggcggcag | aaagaatccc | 5220 |
|----|------|--------------------------|------------------|-------------------|-------------------|------------------------|------------------|--------------|
| | 1007 | cqqattactq | gctggaatac | ggtaacccgt | gggagttcga | gcgccataat | acgcgctaca | 5280 |
| | | | | | | | gtggagaccg | |
| | | | | | | | gacgccacca | |
| | | | | | | taacctcggt | | 5460 |
| | 1011 | agggcgacta | cttcgcggcg | gtggaagata | aaaaccattc | cgagaacgtg | tcgcgggtac | 5520 |
| | 1012 | tctatccgga | tgactcgacc | tattcaggac | gcgagctgcg | cctgcggcag | gagtacttcc | 5580 |
| | 1013 | tegtttegge | gacggtgcag | gacatcctca | gccgccacta | ccagctgcat | aaaacctacg | 5640 |
| | | | | | | | ctgtcgattc | |
| W> | | | | | | | ggnnnnnnn | |
| | | | • | , - | • | | nnnnnnnnn | |
| | | | | | | nnnnnnnnn | | 5880 |
| | | | | | | nnnnnnnnn | | 5940 |
| W> | 1019 | nnnnnnnn | nnnnnnnnn | nnnnnnnnn | nnnnnnnn | nnnnnnnnn | ${\tt nnnnnnnn}$ | 6000 |
| W> | 1020 | nnnnnnnn | nnnnnnnnn | nnnnnnnn | nnnnnnnn | ${\tt nnnnnnnn}$ | ${\tt nnnnnnnn}$ | 6060 |
| W> | 1021 | nnnnnnnn | nnnnnnnn | nnnnnnnn | nnnnnnnn | ${\tt nnnnnnnnn}$ | ${\tt nnnnnnnn}$ | 6120 |
| W> | 1022 | nnnnnnnn | nnnnnnnn | nnnnnnnnn | nnnnnnnn | ${\tt nnnnnnnn}$ | ${\tt nnnnnnnn}$ | 6180 |
| W> | 1023 | nnnnnnnn | nnnnnnnn | ${\tt nnnnnnnn}$ | ${\tt nnnnnnnnn}$ | ${\tt nnnnnnnnn}$ | ${\tt nnnnnnnn}$ | 6240 |
| M> | 1024 | nnnnnnnn | nnnnnnnn | ${\tt nnnnnnnnn}$ | ${\tt nnnnnnnn}$ | ${\tt nnnnnnnn}$ | ${\tt nnnnnnnn}$ | 6300 |
| W> | 1025 | ${\color{blue} nnnnnnn}$ | ${\bf nnnnnnnn}$ | ${\tt nnnnnnnnn}$ | nnnnnnnn | nnnnnnnn | nnnnnnnnn | 6360 |
| | | | | | | nnnnnnnn | | 6420 |
| | | | | | | nnnnnnnnn | | 6480 |
| | | | | | | | nnnnnnnnn | |
| | | | | | | | nnnnnnnn | |
| | | | | | | nnnnnnnn | | 6660 |
| | | | | | | nnnnnnnn | | 6720 |
| | | | | | | | nnnnnnnn | 6780 6840 |
| | | | | | | nnnnnnnn | | 6900 |
| | | | | | | nnnnnnnn | | 6960 |
| | | | | | | nnnnnnnnn | | 7020 |
| | | | | | | nnnnnnnnn nnnnnnnnn | | 7080 |
| | | | | | | nnnnnnnn | | 7140 |
| | | | | | | nnnnnnnn | | 7200 |
| | | | | | | nnnnnnnn | | 7260 |
| | | | | | | nnnnnnnn | | 7320 |
| | | | | | | nnnnnnnn | | 7380 |
| | | | | | | nnnnnnnnn | | 7440 |
| | | | | | | | nnnnnnnn | 7500 |
| | | | | | | nnnnnnnn | | 7560 |
| | | | | | | | ctgcattgat | 7620 |
| | 1047 | ggactgaaag | gtttcgacgg | cggttgtgtg | ggttttgctt | gcccatctgg | cggcatgaat | 7680 |
| | 1048 | agtgtcattc | atgacgatcc | agttcgatat | tcaacagacc | gtctttgtaa | tcggcaccga | 7740 |
| | 1049 | caattttgat | caataaagcg | tttgacctga | tgcatgaggg | taaatccatt | cgttcggttg | 7800 |
| | | | | | | | caatgagtag | |
| | 1051 | tattctgacg | catctgacaa | ttggttccaa | tgacctgaag | aaggcgcgca | tcttttatga | 7920 |
| | 1052 | tgctgttttg | gaaccgttgg | gtatcaaact | tattcgcgag | gtcgaaggac | agcgttttgc | 7980 |
| | 1053 | ctatggtaaa | gacggcgaag | aaggacgcat | catcattgta | aagcctatta | atggtgaagc | 8040 |
| | 1054 | cgctaccgct | ggaaatggta | tcactatcgg | tttggcagcg | ccttctgatg | aagctgtcga | 8100 |
| | | | | | | | | |

RAW SEQUENCE LISTING DATE: 07/05/2001
PATENT APPLICATION: US/09/885,297 TIME: 16:29:04

Input Set : A:\seqlist24CP.txt

Output Set: N:\CRF3\07032001\1885297.raw

1055 tgctttttat aaagcaggct tggctaatgg cggtaaggat gccggagaac cggggcctcg 8160 1056 tccggctgct aataattctc ggggtgccta tttatatgac cctgaaggca ataaaatctg 8220 1057 cgctttcaat tttaaataag atttctttgg tgcagggtta ttcaaaatag ccctgcattt 8280 1058 tcagtattat agcggccatt atggcttttg ccttgataaa aaatttatca gggctgtttt 8340 1059 tcgtgatgaa tatttttgat ttttcaagaa aagcctgata tcttccaaca tctttctttg 8400 1060 tatataaatg gagcgagcta tggcgcgcgt aactgtcgaa gactgtatcg ataaagttca 8460 1061 taatcgtttc gatttgatcc tctagagtca acctgcttgt tactcgtgat cccattcaca 8520 1062 agggcgaatt aattcgccct tctgttccgt taccaacact tgagccggag gcataatggg 8580 1063 aaagccaatg tggcgttgtt gggcgttgat gctgatggtg tggttcagtg cgtcggctac 8640 1064 ggcggcgaac ggctgggaaa tctataaaag ccgtttcatg accacggacg ggcgcattca 8700 1065 ggataccggc aataagaatg tcagccacac cgaaggtcag ggattcgcca tgctgatggc 8760 1066 ggtgcattac gatgaccgca tcgcgttcga taacctgtgg aactggacgc aaagccacct 8820 1067 gcggaacacg accagcggct tgttctactg gcgttacgat ccgtcggcgg ccaatccggt 8880 1068 ggtggataag aacaacgcct cggatggcga tgtgctgatt gcctgggcgt tgttaaaagc 8940 1069 gggaaataag tggcaggaca accgttacct gcaggcgtcg gacagcatcc agaaagcgat 9000 1070 categorage aatateatte agtttgeggg eegeacegtg atgttgeeeg gegeetatgg 9060 1071 tttcaacaag aacagctatg tgatccttaa cccgtcgtat ttcctgttcc cggcctggcg 9120 1072 cgactttgct aaccgcagcc atcttcaggt gtggcggcaa ctgattgacg acagcctgtc 9180 1073 attggtcgga gaaatgcgtt tcggtcaggt cgggctgccg acggactggg cggcgctgaa 9240 1074 cgcggatggc tcgatggcgc cggcgacggc ctggccgtcg cgtttcagtt acgacgccat 9300 1075 tcgtatcccg ctgtatttgt actggtatga cgccaaaacc acggcgctgg tgccgttcca 9360 1076 gctgtactgg cgtaactatc cccgcctgac gacgccggcc tgggttgatg tgctgagcag 9420 1077 taacaccgcg acttacaata tgcagggcgg tttgctggcg gtgcgcgacc tgacgatggg 9480 1078 caacctcgac gggctcagcg atctgccagg cgcatcggaa gattactact cgtcgagcct 9540 1079 gcgcctgctg gtgatgttgg cgcgcggtaa ataaccttat tcttgcggta cacatggcga 9600 1080 ggacgatgtc cttgccattt tccccacttt tatccctctg aatggcgtgt ttttcacgct 9660 1081 ttgttaacct gcttgttact cgtgatccca ttcacaaggg cgaattgacc tgcaggcatg 9720 1082 caagettgge gtaatcatgg teatagetgt tteetgtgtg aaattgttat eegeteacaa 9780 1083 ttccacacaa catacgagcc ggaagcataa agtgtaaagc ctggggtgcc taatgagtga 9840 1084 gctaactcac attaattgcg ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt 9900 1085 gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgct 9960 1086 cttccgctgg cgcgccaggt cgactctaga ggatccccgg ggaagatctt ccggaagatc 10020 1087 ttcccgagct cgaattaatt ccgcgatgaa ttgatcccgg aagttcctat tctctagaaa 10080 1088 gtataggaac tcgaattggt cgacaagcta gcttgcatgc aagcttgtat tctatagtgt 10140 1089 cacctaaatc gtatgtgtat gatacataag gttatgtatt aattgtagcc gcgttctaac 10200 1090 gacaatatgt acaagcctaa ttgtgtagca tctggcttac tgaagcagac cctatcatct 10260 1091 ctctcgtaaa ctgccgtcag agtcggtttg gttggacgaa ccttctgagt ttctggtaac 10320 1092 gccgttccgc accccggaaa tggtcagcga accaatcagc agggtcatcg ctagcccatg 10380 1093 gctaattctg tcagccgtta agtgttcctg tgtcactgaa aattgctttg agaggctcta 10440 1094 agggettete agtgegttae atecetgget tgttgteeae aacegttaaa eettaaaage 10500 1095 tttaaaagcc ttatatattc tttttttct tataaaactt aaaaccttag aggctattta 10560 1096 agttgctgat ttatattaat tttattgttc aaacatgaga gcttagtacg tgaaacatga 10620 1097 gagettagta egttageeat gagagettag taegttagee atgagggttt agttegttaa 10680 1098 acatgagagc ttagtacgtt aaacatgaga gcttagtacg tgaaacatga gagcttagta 10740 1099 cgtactatca acaggttgaa ctgcggatct tgcggccgca aaaattaaaa atgaagtttt 10800 1100 gacggtatcg aaccccagag tcccgctcag aagaactcgt caagaaggcg atagaaggcg 10860 1101 atgcgctgcg aatcgggagc ggcgataccg taaagcacga ggaagcggtc agcccattcg 10920 1102 ccgccaaget etteageaat ateaegggta gecaaegeta tgteetgata geggteegee 10980 1103 acacccagcc ggccacagtc gatgaatcca gaaaagcggc cattttccac catgatattc 11040

RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION:

US/09/885,297

TIME: 16:29:04

Input Set : A:\seqlist24CP.txt

Output Set: N:\CRF3\07032001\I885297.raw

1104 ggcaagcagg categceatg ggtcacgacg agatectege cgtegggeat cegegeettg 11100
1105 ageetggega acagttegge tggegegage ceetgatget ettegtecag ateatectga 11160
1106 tegacaagae eggettecat eegagtaegt getegetega tgegatgttt egettggtgg 11220
1107 tegaatggge aggtageegg ateaagegta tgeagegee geattgeate ageeatgatg 11280
1108 gatacttee eggeaggage aaggtgagat gacaaggagat eetgeeeegg eacttegeee 11340
1109 aatageagee agteeettee egetteagtg acaaegtega geaeagetge geaaggaaeg 11400
1110 eeegtegtgg eeageeaega tageegeget geetegtett gaggteatt eagggeaeeg 11460
1111 gacaggtegg tettgacaaa aagaaeeggg egeeeetgeg etgaeageeg gaaeaeggeg 11520
1112 geateagage ageegattg etgttgee eagteatage egaatageet etecaeeeaa 11580
1113 geggeeggag aacetgegtg eaateeatet tgtteaatea tgegaaaega teeteateet 11640
1114 gtetettgat eeactagatt attgaageat ttateagggt tattgtetea tgageggata 11700
1115 eatatttgaa tgtatttaga aaaataaaea aataggggtt eegegeaeat tteeeegaaa 11760
1116 agtgeeaeet ge